The basis of musical consonance as revealed by congenital amusia

Marion Cousineau[a,1], Josh H. McDermott[b], and Isabelle Peretz[a]

[a]International Laboratory for Brain, Music and Sound Research (BRAMS) and Center for Research on Brain, Language and Music (CRBLM), Department of Psychology, University of Montreal, Montreal, QC, Canada H3C 3J7; and [b]Center for Neural Science, New York University, New York, NY 10003-6603

Edited by Dale Purves, Duke–National University of Singapore Graduate Medical School, Republic of Singapore, and approved October 9, 2012 (received for review May 11, 2012)

Some combinations of musical notes sound pleasing and are termed "consonant," but others sound unpleasant and are termed "dissonant." The distinction between consonance and dissonance plays a central role in Western music, and its origins have posed one of the oldest and most debated problems in perception. In modern times, dissonance has been widely believed to be the product of "beating": interference between frequency components in the cochlea that has been believed to be more pronounced in dissonant than consonant sounds. However, harmonic frequency relations, a higher-order sound attribute closely related to pitch perception, has also been proposed to account for consonance. To tease apart theories of musical consonance, we tested sound preferences in individuals with congenital amusia, a neurogenetic disorder characterized by abnormal pitch perception. We assessed amusics' preferences for musical chords as well as for the isolated acoustic properties of beating and harmonicity. In contrast to control subjects, amusic listeners showed no preference for consonance, rating the pleasantness of consonant chords no higher than that of dissonant chords. Amusics also failed to exhibit the normally observed preference for harmonic over inharmonic tones, nor could they discriminate such tones from each other. Despite these abnormalities, amusics exhibited normal preferences and discrimination for stimuli with and without beating. This dissociation indicates that, contrary to classic theories, beating is unlikely to underlie consonance. Our results instead suggest the need to integrate harmonicity as a foundation of music preferences, and illustrate how amusia may be used to investigate normal auditory function.

music perception | aesthetic preferences | roughness

Music is made by combining notes, but not all combinations are comparable aesthetically. When notes are produced simultaneously, some combinations sound pleasing and are termed "consonant," and others sound unpleasant or out of place, and are termed "dissonant" (1–3). This contrast forms one of the key ingredients of music composition, in which dissonant chords are used to create feelings of tension that are later released by consonant chords. Although their perception is in some cases dependent on context (4), marked differences between consonance and dissonance are typically apparent even in isolated chords (5–8). Preferences for consonance appear to be present in human infants (9–12) and perhaps also in other species with little exposure to music (13, 14, although see also ref. 15), consistent with a biological basis. However, the possibility also remains that consonance preferences are learned—such preferences are correlated with musical experience in Western listeners (8), and their existence in non-Western cultures is controversial (16–18). Consonance has fascinated scholars since the time of the Greeks, in part because consonant and dissonant chords can be viewed as atomic aesthetic elements, the investigation of which provides a point of entry to understanding the roles of biology and culture in music perception.

Greek scholars believed that consonance was related to the ratio between the length of the strings on which notes were sounded, inspired by ratio-based theories of aesthetics that applied to domains ranging from music to architecture. However, contemporary thinking on consonance is instead rooted in acoustics, beginning with the fact that musical instrument and voice sounds are composed of multiple discrete frequencies. These frequencies are termed "harmonics" because they are typically integer multiples of the fundamental frequency (F0, inverse of the period) of the sound. Harmonics are combined in a single waveform when traveling in the air but are partly segregated by the cochlea, because different auditory nerve fibers respond to different frequencies (19).

When several notes are combined to form a chord, the resulting sound waveform that enters the ear contains all of the individual frequencies of each note. Auditory scientists have long noted that aspects of the pattern of component frequencies differ between consonant and dissonant chords. Prevailing theories ascribe consonance to the fact that dissonant chords contain frequency components that are too closely spaced to be resolved by the cochlea (5, 6, 20–22). Two such components shift in and out of phase over time, producing an interaction that oscillates between constructive and destructive interference. The amplitude of the combined physical waveform thus alternately waxes and wanes. If the components are close enough to excite the same set of auditory fibers, amplitude modulations are directly observable in the response of the auditory nerve (23). These amplitude modulations are called "beats," and result in an unpleasant sensation known as "roughness," analogous to the tactile roughness felt when touching a corrugated surface [in practice, the perception of roughness is dependent on the depth and rate of amplitude modulation, as well as the center frequency of the tones involved (24)]. Beats are mostly inaudible when two pure tones are presented independently to separate ears, presumably because they do not interact at the level of the cochlea [although "binaural beats" can be audible for very small frequency differences (25)]. Helmholtz (20) is usually credited with the idea that dissonant chords are unpleasant because they contain the sensation of roughness, a notion that was fleshed out by ensuing generations of psychoacousticians (5, 6, 21, 22, 26). Theories of dissonance based on beating have been dominant in the last century and are now a regular presence in textbooks (1–3).

However, a second acoustic property also differentiates consonance and dissonance: the component frequencies of the notes of consonant chords combine to produce an aggregate spectrum that is typically harmonic, resembling the spectrum of a single sound with a lower pitch (Fig. 1*B*). In contrast, dissonant chords produce an inharmonic spectrum. Such observations led to a series of analyses and models of consonance based on harmonicity (23, 27–29). Although beating-based theories are widely accepted as the standard account of consonance, harmonicity has remained a plausible alternative.

One reason these candidate explanations of consonance have proven difficult to test is that they share qualitative predictions, because dissonant chords both contain inharmonic frequencies

Author contributions: M.C., J.H.M., and I.P. designed research; M.C. performed research; M.C. analyzed data; and M.C., J.H.M., and I.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: marioncousineau@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1207989109/-/DCSupplemental.

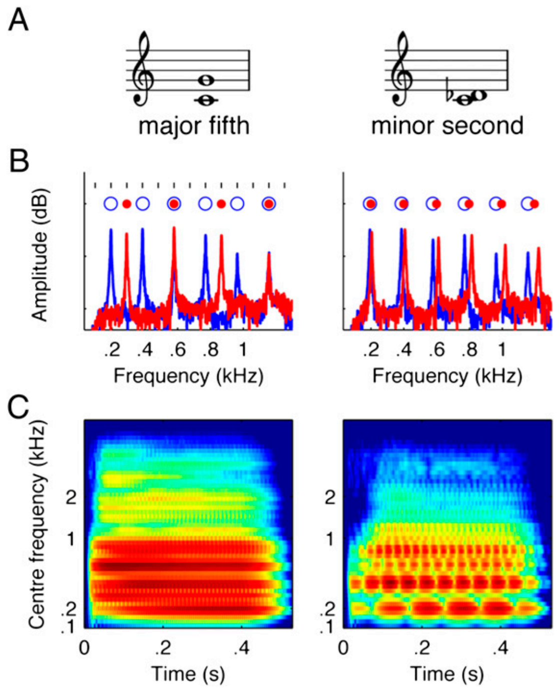

Fig. 1. Example chords: the perfect fifth (*Left*; conventionally consonant), and the minor second (*Right*; conventionally dissonant) composed of saxophone notes. (*A*) Musical notation. (*B*) Amplitude spectra. Spectrum of lower (root) note is shown in blue, that of higher note in red. Blue and red dots denote the frequencies belonging to each note. In the consonant perfect fifth, the two notes share common frequencies and produce an aggregate spectrum that is harmonic (corresponding to a harmonic series with a lower F0, indicated by the vertical bars at the top). In the dissonant minor second, the notes share no harmonics and the overall spectrum is inharmonic. (*C*) Predicted activation of auditory nerve fibers (55) for each chord. The beating produced by the closely spaced frequencies of the minor second is visible in the low frequency channels of the right-hand plot.

and produce beats. However, it is possible to produce synthetic sounds that isolate the two acoustic properties. A recent study (8) used such stimuli to examine individual differences in acoustic preferences within large cohorts of subjects. It was found that the degree of preference for consonance in individual subjects correlated with the degree of preference for harmonicity but not with the degree of aversion to beating. These results support a theory of consonance based on harmonicity but are limited by the restricted variance in behavioral responses obtained with populations of normal listeners. In particular, nearly all normal listeners possess some degree of preference for consonance and aversion to beating, leaving open the possibility that beating contributes in an important way to their perception of consonance.

Here, we adopted a neuropsychological approach to investigate the basis of consonance by testing individuals with congenital amusia. Congenital amusia is a neurogenetic disorder that is characterized by a deficit in melody processing that cannot be explained by hearing loss or intellectual deficiencies (30). Congenital amusia is hereditary (31, 32) and is associated with reduced connectivity between the right auditory and inferior frontal cortices (33–35). The root functional cause of the musical impairments appears to lie in the processing of pitch. Amusic individuals are impaired at detecting the direction of pitch changes between notes (36, 37) as well as pitch deviations (irrespective of direction) that are small (e.g., less than two semitones) (38, 39). As pitch changes in typical music are often below their abnormally high thresholds (40), amusics are bound to miss an essential part of musical structure.

The aim of the present study was to exploit the perceptual deficits observed in congenital amusia to dissect the basis of consonance. There is one previous demonstration that amusics do not show the normally observed preference for consonant over dissonant music (30). In that study, amusics were presented with classical musical excerpts containing melodies and accompaniment on the piano. The amusics were asked to rate the pleasantness of the original excerpts as well as manipulated versions in which the accompaniment was pitch-shifted to introduce dissonance. Amusics gave similar ratings to the two versions, suggesting deafness to dissonance. We sought to understand the basis of this phenomenon. We first assessed the perception of consonant and dissonant chords in 10 amusic and 10 matched control subjects. Synthetic stimuli were then used to investigate the perception of harmonicity and beating.

Results

Consonant and Dissonant Chords. In the first experiment, participants were asked to rate the pleasantness of chords (intervals or triads composed of notes sung by a vocalist or played on a saxophone) on a nine-step scale. The chords included the unison, the intervals spanning the minor second (1 semitone) to the major seventh (11 semitones), and the four most common triads in Western music. Fig. 2 shows the pleasantness ratings of the two subject groups. Control subjects' ratings were consistent with previous results (8), with conventionally dissonant chords such as the minor second or major seventh rated as unpleasant, and conventionally consonant chords such as the major third or major triad rated as pleasant. In contrast, amusics' ratings varied across subjects, generally with little resemblance to the pattern of results in normal listeners. On

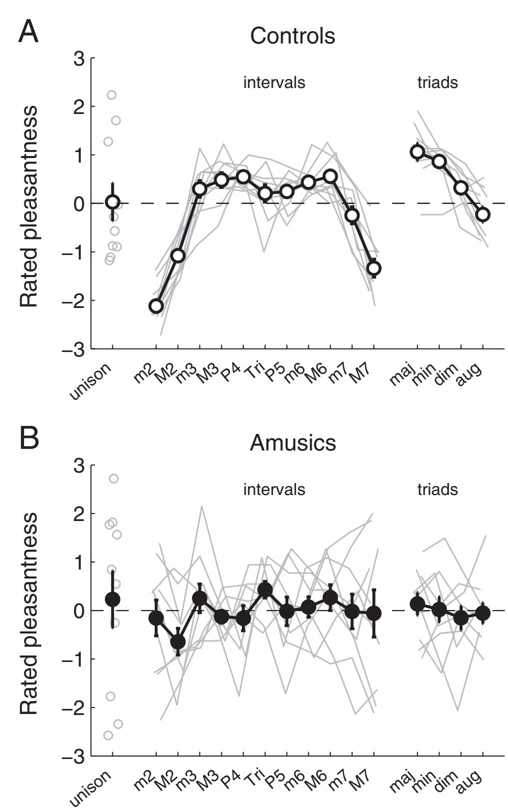

Fig. 2. Mean (black) and individual (gray) pleasantness ratings of chords by the control (*A*) and amusic (*B*) groups. Ratings of each subject were *z*-scored before analysis to remove coarse differences in the use of the rating scale, as in ref. 8. Intervals contained two notes separated by an integer number of semitones; triads (major, minor, diminished, and augmented) contained three notes. Here and elsewhere, error bars denote one SE.

average, amusics gave similar ratings to all chords, with consonant chords rated no more highly than dissonant chords.

To assess agreement among raters in the two populations, we estimated the null distribution of Kendall's W for ratings of 16 conditions by 10 subjects. Ninety-five percent of 10,000 W values obtained from randomly drawn ratings were situated below 0.164. Control subjects' ratings yielded a W of 0.68, showing far more consistency than would be expected by chance. In contrast, amusics' ratings were inconsistent (W = 0.08), no different from what would be expected by chance.

We note that the ratings obtained from controls, despite generally resembling normal ratings, show less-pronounced rating variation for the medium-sized intervals, particularly the tritone, than what was reported in ref. 8, for example, with identical stimuli. This result might reflect two characteristics of our control population (matched to the demographics of the amusic group as closely as possible) that are known to weaken the strength of consonance preferences: some subjects showed mild to moderate hearing loss (41), and all subjects had very little, if any musical training (8). Overall, however, the results of the control group are largely as expected given prior work.

To test whether any amusic showed a pattern of results close to that of the controls, we computed correlations between each individual participant's ratings of the 16 chord types and the mean ratings of the control group. As shown in Fig. 3*A*, the correlations between the amusics' data and the average ratings of the control group were small, never exceeding that expected by chance, whereas those of each control subject (correlated with the mean of the ratings obtained from the nine other controls) were statistically significant.

To test whether the participants exhibited internally consistent rating patterns, we measured test-retest reliability for individual subjects, using correlations between ratings averaged over the first and second half of the trials for each chord type. As shown in Fig. 3*B*, all control subjects gave consistent ratings, whereas only two of the amusics did. These two subjects' ratings seem to be based on idiosyncratic pitch height preferences, rather than on consonance (one subject preferred high-pitched tones, and the other preferred low-pitched tones, yielding preferences that covaried with interval size). The amusics thus never resembled control subjects with respect to consonance judgments, and were inconsistent in their ratings, indicating that consonance and dissonance as heard by normal listeners played little role in their perception of chords.

Pleasantness of Emotional Vocal Stimuli. To ensure that the group differences in chord ratings were not a result of general differences in the ability to ascribe pleasantness ratings to sounds, we measured pleasantness ratings of emotional vocalizations [Montreal Affective Voices (42)]. The stimuli were recordings of nonverbal vocal sounds made by humans to express sadness, fear, surprise, or happiness (cries, gasps, laughing, etc.). In contrast to the results with chords, the two groups showed nearly identical rating patterns for vocalizations (Fig. 4), rating happy vocalizations the highest and sad and fearful vocalizations the lowest. This result indicates that the groups made similar use of the rating scale, and that amusics do not differ generally from controls in the way they assign pleasantness ratings to sounds.

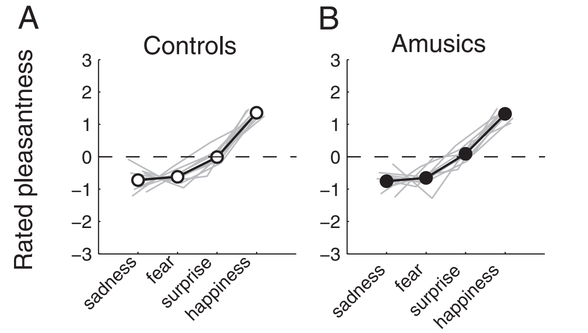

Fig. 4. Mean (black) and individual (gray) pleasantness ratings of emotional vocalizations by control (*A*) and amusic (*B*) groups. Ratings of each subject were *z*-scored. Error bars are not visible as they were smaller than the symbol size.

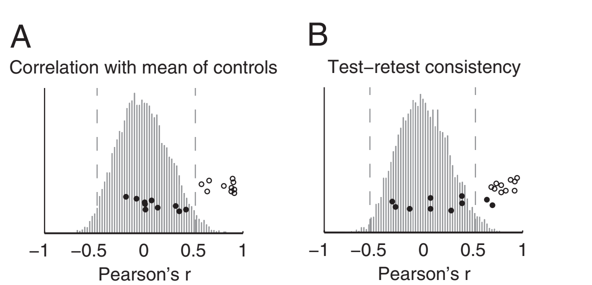

Fig. 3. (*A*) Correlations between individual ratings of chords and mean chord ratings of the control group. Null distribution is drawn from correlations between random permutations of individual control subject's data and the mean of the control group. Vertical dashed lines here and in *B* denote 95% confidence intervals around zero. (*B*) Test-retest correlations between ratings averaged over first and second halves of the experiment for each subject. Null distribution is drawn from correlations between the first half and a randomly permuted second half of randomly selected individual control subject's data.

Harmonicity and Beating. In a third experiment, we investigated preferences for harmonicity and beating using stimuli taken from a previous study in normal listeners (8). Preference measures were derived from differences of ratings between stimuli that either did or did not contain these acoustic factors (harmonic and inharmonic, nonbeating and beating), but that were matched in other respects (number of frequencies, amplitude, etc.).

Harmonicity was isolated by comparing ratings of harmonic complex tones ("harmonic-subset" and "octave-spaced") with those of inharmonic complex tones generated by subtly perturbing the frequencies of the harmonic tones (Fig. 5*A*: Jittered, Shifted, and 13st-spaced). All tones contained widely spaced frequency components to avoid beating. The difference between the average ratings of harmonic and inharmonic tones provided "measure 1" of harmonicity preference (Fig. 5*C*). We also computed a second measure of harmonicity preference from the difference of the ratings for pure tones (the simplest harmonic stimulus) and dichotically presented tone pairs (which were inharmonic by virtue of being closely spaced, but produced little perceptible beating because of the absence of interactions at the level of the cochlea) (8); this was termed "measure 2" of harmonicity preference (Fig. 5*C*). Beating was isolated using the fact that dichotic presentation of two tones (one tone to the left ear and the other to the right ear, via headphones) prevents their interaction at the level of the cochlea, and thus mostly eliminates the perception of beating (25). We contrasted ratings of pairs of closely spaced pure tones presented either diotically (both tones to both ears, producing substantial beating) or dichotically (minimizing beating) (Fig. 5*B*). For comparison of these harmonicity and beating preferences with consonance preferences, we computed analogous measures for the preference for consonant dyads (e.g., perfect fifth) over dissonant dyads (e.g., minor second) and for consonant triads (major triad) over dissonant triads (augmented triad) using the ratings from Exp. 1.

Control subjects showed preferences for consonance, harmonicity, and the absence of beats, as expected (Fig. 5*C*). In contrast, amusic listeners failed to show preferences for consonance or harmonicity but exhibited an aversion to beating comparable to that of controls. The two populations' preferences were significantly different for both consonance and both harmonicity measures ($P < 0.001$ in all cases; *t* tests) but not for beating ($P = 0.71$).

Harmonicity and Beating Discrimination. In the fourth experiment, we assessed whether the lack of preference for harmonicity in amusics was because of an inability to discriminate harmonic from inharmonic frequency relations. On each trial, subjects were presented with two harmonic tones and one of its inharmonic counterparts (generated by jittering or shifting the spectrum of the harmonic tones). The three stimuli were generated with distinct F0s, such that the discrimination had to be made on the basis of the frequency relations within the tones rather than any individual frequency. Subjects were asked to report which of the three stimuli sounded as if played by a different instrument. Whereas control subjects scored high on this task, amusics performed poorly ($P < 0.00001$; t test), suggesting that harmonic and inharmonic spectra are poorly distinguished by their auditory system (Fig. 6).

For comparison, we also tested discrimination between our beating and nonbeating stimuli (Exp. 5). On each trial, subjects were presented with two dichotic (nonbeating) and one diotic (beating) sounds, again with distinct F0s. Subjects were again asked to report which of the three stimuli sounded as if played by a different instrument. Performance was high in both groups (Fig. 6) and did not differ significantly between groups ($P = 0.53$; t test), suggesting normal perception of beating in amusics despite abnormal harmonicity perception.

Discussion

We tested the perception of dissonance, harmonicity, and beating in amusic listeners. The pleasantness ratings for chords showed that amusics have highly abnormal perception of consonance and dissonance. Whereas control participants preferred consonance to dissonance, amusic participants were indifferent to the distinction. Amusic listeners' perception of harmonicity was also impaired. Unlike controls, amusics did not prefer harmonic to inharmonic tones and could not discriminate them from each other. In contrast to these marked deviations from normal perception, amusics' judgments of beating were indistinguishable from those of normal listeners; every amusic participant preferred stimuli without beating over those with beats, to a degree comparable to that of control subjects, and every amusic could distinguish sounds with and without beating.

Implications for the Basis of Consonance and Dissonance. The unique contribution of this study to the understanding of consonance was to use congenital amusia to demonstrate a dissociation between the

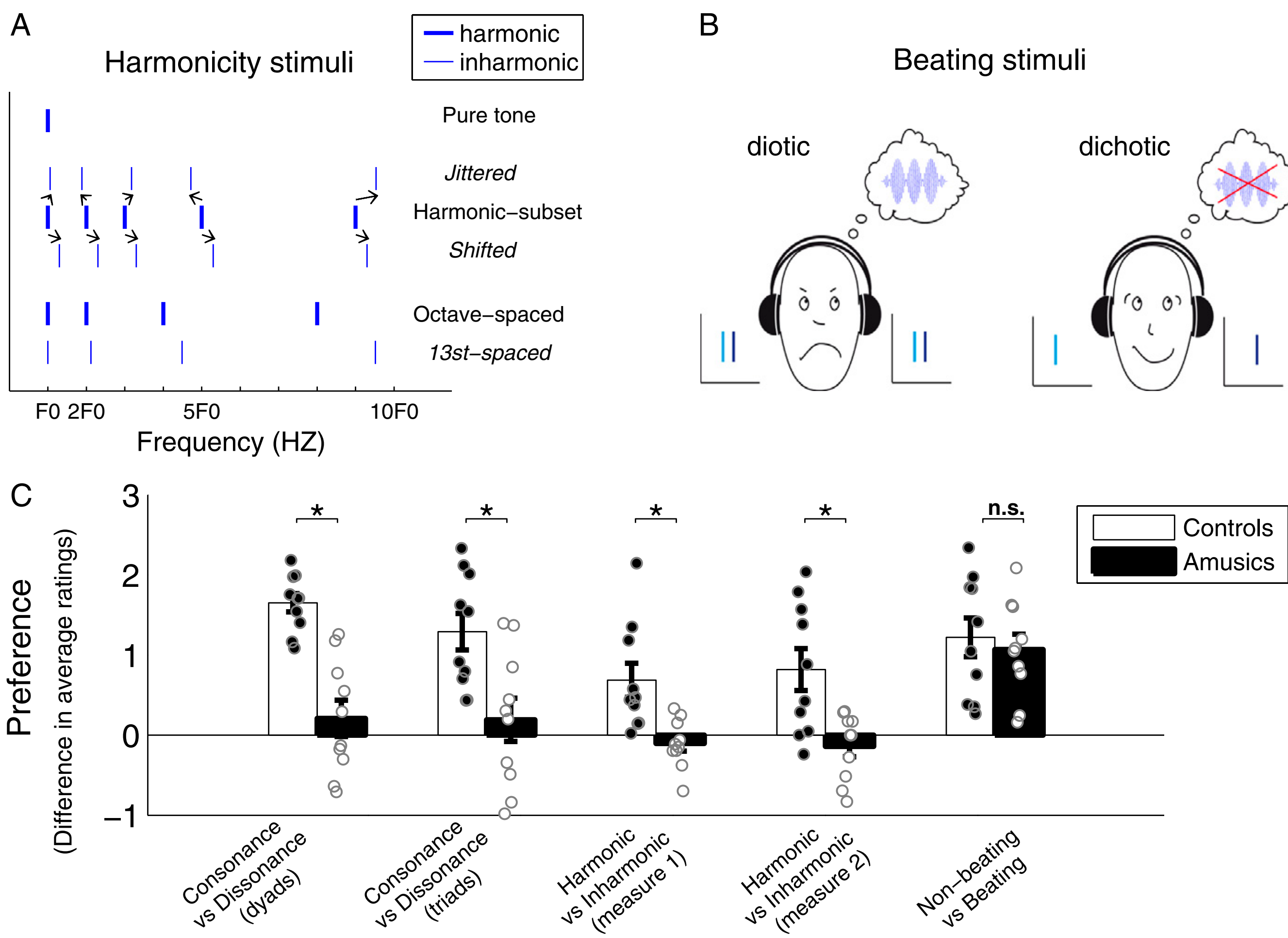

Fig. 5. (*A*) Schematic spectra of harmonicity test stimuli (Exps. 3 and 4): Thick and thin lines depict spectra of harmonic and inharmonic tones, respectively. See *Materials and Methods* for details. (*B*) Beating stimuli (Exps. 3 and 5) were made of two pure tones close in frequency (0.75 or 1.5 semitones apart). They were presented diotically (both tones to both ears), resulting in the perception of beats, or dichotically (one tone to each ear), producing a substantial decrease in perceived beating. (*C*) Mean and individual preferences for the two groups of listeners. Preference measures, as in ref. 8, were differences in the average ratings for two groups of sounds. Consonance (dyads): mean rating (MR) for the four most consonant dyads minus MR for the four most dissonant dyads; Consonance (triads): MR for the major triad minus MR for the augmented triad; Harmonicity1: MR for the harmonic complexes minus MR for the inharmonic complexes; Harmonicity2: MR for the pure tones minus MR for the dichotic pairs of pure tones; Beating: MR for the dichotic pairs of pure tones minus MR for the diotic pairs of pure tones.

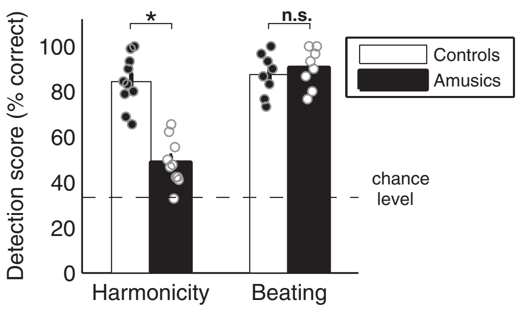

Fig. 6. Mean and individual performance for harmonicity (Exp. 4) and beating (Exp. 5) discrimination tasks. Horizontal dashed line represents chance level in this 3-AFC task: 33.33%.

perception of beating and the perception of dissonance. Amusics exhibit the normal aversion to beating, but lack the normal preference for consonance over dissonance, suggesting that consonance preferences do not derive from an aversion to beating. Rather, our results suggest that the quality of roughness (induced by beating) constitutes an aesthetic dimension that is distinct from that of dissonance. This conclusion is consistent with earlier work in normal listeners, suggesting that harmonicity is more closely related to consonance than is beating (8), but the present study provides the unique demonstration of a clear-cut dissociation, something not possible in normal listeners. Amusia can thus provide a tool for studying the underpinnings of music perception.

Given the prevalence of beating-based accounts of dissonance (5, 6, 21), it is natural to ask why amusics' aversion to beating is not sufficient to produce preferences for consonance over dissonance. We believe the explanation is that for naturally occurring musical chords, beating is not in fact closely associated with dissonance. Naturally occurring tones produced by instruments or voices generally have spectra whose component amplitudes decrease with increasing frequency. As a result, roughness in naturally occurring chords originates primarily from the interaction of the first few frequency components of the notes. In contrast, classic studies of dissonance (e.g., ref. 5) relied on artificial synthetic tones with equal amplitude components, in which beating can be produced by all components in the spectrum, yielding artificially strong and consistent beating. Moreover, roughness is influenced by factors such as the amplitude (43) or the phase (44) of frequency components, both of which can vary greatly from one instrument to another, or with the way keys or strings are sounded on a given instrument. Consistent with these ideas, we have found that for real-world musical notes, as were used in our experiments, beating is not consistently more pronounced in dissonant chords than in consonant chords, that it often varies more with timbre than with chord type, and that overall amounts of beating in real-world chords are often far below that in the synthetic chords used in classic psychoacoustic studies (See Figs. S1–S3 and *SI Text*). It may be that beating is in practice sufficiently low or variable that it has little effect on the average aesthetic response to a chord. Alternatively, the inconsistency of beating across instances of a chord may cause listeners to implicitly learn to ignore it in musical contexts, where it might otherwise cause consonance judgments to fluctuate across instruments.

Our results are consistent with the alternative account of consonance in terms of harmonicity (8, 23, 27–29), in that amusics were consistently abnormal in their perception of both consonance and harmonicity. Unlike the perception of beating, harmonicity is not strongly dependent on frequency amplitudes or phases (45). This relative invariance might explain the similarity in consonance judgments for different instrument timbres (8). Models measuring harmonicity with periodicity detection mechanisms (29) can predict the degree of consonance of musical intervals, and harmonicity preferences in normal listeners have been shown to be correlated with their consonance preferences (8), suggesting a link (see also *SI Text* and Fig. S4 for correlational analyses of the present data). Differential auditory brainstem responses to consonant and dissonant intervals (23, 46) are also consistent with a role for harmonicity in consonance perception.

The origins of the aesthetic response to harmonicity remain to be determined. One might envision that humans are born with preferences for harmonicity, or that they derive from exposure to natural sounds because of the prevalence of harmonic spectra in mammalian vocalizations (47), or that they derive from exposure to music (8). Our results leave each of these possibilities open, but indicate that the nature/nurture debate surrounding consonance should be oriented toward harmonicity rather than beating/roughness.

Dissonance, Pitch, and Amusia. The pitch- and rhythm-related musical abilities of congenital amusic individuals have been extensively studied over the last decade (see ref. 48 for a review). Whereas only some amusics show impaired perception of rhythm, deficits in pitch perception appear to be universally present and have thus been considered the hallmark of amusia (36, 38). Recently, however, it was reported that amusic listeners also have a deficit in memory for pitch (49–51). This memory impairment could potentially play a role in all previously reported amusic deficits, because threshold measurements or melody comparisons require participants to compare sounds presented sequentially, and thus to store pitch-related information over time. The deficits evident in the present study are less vulnerable to this possibility because pleasantness ratings do not require sounds to be held in memory. Our results thus suggest that amusia cannot be solely caused by memory impairments and indeed reflects defective encoding of pitch information.

The deficit we observed in the perception of harmonicity in amusic listeners further suggests that the root of congenital amusia may lie in the inadequate representation of this basic acoustic property of pitched sounds. A harmonicity-based deficit would likely extend outside of music. Amusic deficits in prosody perception remain controversial (30, 37), but our results suggest other domains where amusics and normal listeners may differ. Voice attractiveness, for instance, is positively correlated with the degree of harmonicity (52). Harmonicity is also an important cue for concurrent sound segregation (53). Assessing amusic individuals with tasks involving such nonmusical harmonicity-dependent abilities is an intriguing direction for future research. In particular, our results raise the possibility that amusia could be used to test the role of pitch in these and other aspects of normal auditory function.

Materials and Methods

Subjects and Set-Up. Ten amusic and 10 matched control participants were tested. The two groups did not differ significantly in age, years of education, years of musical training, and audiometric thresholds (Table S1). Each amusic scored 2 SD below the mean of the general population when tested with the Montreal Battery of Evaluation of Amusia, six tests that assess musical pitch perception, rhythm perception, and memory (54). Stimuli were delivered by an external soundcard (RME Fireface 800) with a 48-kHz sampling rate and 16-bit resolution. Sounds were played out with closed headphones (Beyerdynamic DT 770 Pro) at 65 dB SPL. Participants were tested in a sound-insulated booth. All listeners were fully informed about the goal of the present study and provided written consent before their participation.

Stimuli. ***Chords.*** Natural recordings of single notes from the equal-tempered scale (from a saxophone or a trained female vocalist) were summed to create intervals or triads. Eight different versions of each type of chord were used, each with a different root note, ranging from G#3 to D#4.

Harmonicity stimuli. Harmonic and inharmonic complexes were generated by summing pure tones with aligned starting phases. The complexes amplitudes were decreased by 14 dB per octave to resemble natural sound spectra. Harmonic stimuli comprised: (*i*) pure tone, (*ii*) harmonic-subset (F0, 2F0, 3F0, 5F0, and 9F0), and (*iii*) octave-spaced (F0, 2F0, 4F0, 16F0). Inharmonic stimuli comprised: (*i*) jittered (even harmonics of the harmonic-subset tones were shifted up and odd harmonics were shifted down, by half a semitone), (*ii*) shifted (all harmonics of the harmonic-subset tones were shifted up by 30 Hz), and (*iii*) 13st-spaced (four frequency components, each 13 semitones apart). We also tested dichotic versions of these stimuli (see *SI Text* and Figs. S5 and S6 for motivation and results).

Beating stimuli. Stimuli consisted of pairs of pure tones, either 0.75 or 1.5 semitones apart. These frequency separations are small enough that both tones fall within the same critical band when presented diotically, and thus produce a fair amount of beating. Nonbeating stimuli consisted of the same pure tones presented dichotically, largely eliminating perceived beating (8, 25).

The stimuli of Exps. 1, 3, 4, and 5 (chord, harmonicity, and beating test stimuli) were taken from a previous study (8). See Audios S1, S2, S3, S4, S5, S6, S7, S8, S9, and S10 for audio examples of stimuli.

Procedure. Responses were collected using a keyboard. Sounds were presented automatically 500 ms after the subject's previous response.

Pleasantness ratings. Chords of Exp. 1 were presented in two blocks of 128 trials each, in which all 16 chord types and eight root pitches were randomly interleaved. Saxophone and voice chords were presented in separate blocks. Vocalizations of Exp. 2 were presented in a single block of 80 trials in which all 10 speakers and four types of emotions were repeated once and randomly interleaved. Exp. 3 was divided into two blocks of 69 trials that featured all types of harmonicity and beating stimuli randomly interleaved. The F0s of the stimuli were randomly assigned to C#4, D#4, or F4 in one of the blocks and to D4, E4, or F#4 in the other block.

Discrimination experiments. Exp. 4 was carried out in one block of 90 trials. On each trial, one randomly chosen interval contained the inharmonic tone; the two others contained its harmonic counterpart (harmonic-subset for shifted and jittered, and octave-spaced for 13st-spaced). Trial types were randomly intermixed within the block. The three tones had F0s two semitones apart (C#4, D#4, F4) or (D4, E4, F#4) in a random order. Exp. 5 was carried out in one block of 30 trials in which the 0.75 and 1.25 semitones pairs of pure tones were randomly interleaved. On each trial, one randomly chosen interval contained a pair of tones presented diotically; the other two contained pairs of tones with the same frequency separation presented dichotically. The most salient difference between the two types of stimuli was thus the presence or absence of beating. The lower tones of the stimuli within a trial were either (C#4, D#4, F4) or (D4, E4, F#4), randomly assigned within and across trials. Feedback was provided in both discrimination experiments.

ACKNOWLEDGMENTS. This work was supported by a postdoctoral grant from the Fyssen Foundation (to M.C.); the Howard Hughes Medical Institute (to J.H.M.); and grants from Natural Sciences and Engineering Research Council of Canada, the Canadian Institutes of Health Research, and from the Canada Research Chairs program (to I.P.).

1. Dowling WJ, Harwood DL (1986) *Music Cognition* (Academic, Orlando, FL).
2. Krumhansl CL (1990) *Cognitive Foundations of Musical Pitch* (Oxford Univ Press, New York).
3. Deutsch D (2009) *The Psychology of Music* (Academic, San Diego), 2nd Ed.
4. Johnson-Laird PN, Kang OE, Leong YC (2012) On musical dissonance. *Music Percept* 30(1):19–35.
5. Plomp R, Levelt WJ (1965) Tonal consonance and critical bandwidth. *J Acoust Soc Am* 38(4):548–560.
6. Kameoka A, Kuriyagawa M (1969) Consonance theory part I: Consonance of dyads. *J Acoust Soc Am* 45(6):1451–1459.
7. Kameoka A, Kuriyagawa M (1969) Consonance theory part II: Consonance of complex tones and its calculation method. *J Acoust Soc Am* 45(6):1460–1469.
8. McDermott JH, Lehr AJ, Oxenham AJ (2010) Individual differences reveal the basis of consonance. *Curr Biol* 20(11):1035–1041.
9. Trainor LJ, Tsang CD, Cheung VHW (2002) Preference for sensory consonance in 2- and 4-month-old infants. *Music Percept* 20(2):187–194.
10. Zentner MR, Kagan J (1996) Perception of music by infants. *Nature* 383(6595):29.
11. Masataka N (2006) Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. *Dev Sci* 9(1):46–50.
12. Perani D, et al. (2010) Functional specializations for music processing in the human newborn brain. *Proc Natl Acad Sci USA* 107(10):4758–4763.
13. Sugimoto T, et al. (2010) Preference for consonant music over dissonant music by an infant chimpanzee. *Primates* 51(1):7–12.
14. Chiandetti C, Vallortigara G (2011) Chicks like consonant music. *Psychol Sci* 22(10): 1270–1273.
15. McDermott J, Hauser M (2004) Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. *Cognition* 94(2):B11–B21.
16. Butler JW, Daston PG (1968) Musical consonance as musical preference: A cross-cultural study. *J Gen Psychol* 79(1st Half):129–142.
17. Maher TF (1976) "Need for resolution" ratings for harmonic musical intervals: A comparison between Indians and Canadians. *J Cross Cult Psychol* 7(3):259–276.
18. Fritz T, et al. (2009) Universal recognition of three basic emotions in music. *Curr Biol* 19(7):573–576.
19. Plomp R (1964) The ear as a frequency analyzer. *J Acoust Soc Am* 36:1628–1636.
20. Helmholtz H (1954) *On the Sensation of Tone as a Physiological Basis for the Theory of Music* (Dover, New York).
21. Sethares WA (1993) Local consonance and the relationship between timbre and scale. *J Acoust Soc Am* 94(3):1218–1228.
22. Hutchinson W, Knopoff L (1978) The acoustic component of western consonance. *J New Music Res* 7(1):1–29.
23. Tramo MJ, Cariani PA, Delgutte B, Braida LD (2001) Neurobiological foundations for the theory of harmony in western tonal music. *Ann N Y Acad Sci* 930:92–116.
24. Terhardt E (1974) On the perception of periodic sound fluctuations (roughness). *Acustica* 30(4):202–213.
25. Rutschmann J, Rubinstein L (1965) Binaural beats and binaural amplitude-modulated tones: Successive comparison of loudness fluctuations. *J Acoust Soc Am* 38(5):759–768.
26. Vassilakis P (2005) Auditory roughness as a means of musical expression. *Selected Reports in Ethnomusicology* 12:119–144.
27. DeWitt LA, Crowder RG (1987) Tonal fusion of consonant musical intervals: The oomph in Stumpf. *Percept Psychophys* 41(1):73–84.
28. Huron D (1991) Tonal consonance versus tonal fusion in polyphonic sonorities. *Music Percept* 9(2):135–154.
29. Ebeling M (2008) Neuronal periodicity detection as a basis for the perception of consonance: A mathematical model of tonal fusion. *J Acoust Soc Am* 124(4): 2320–2329.
30. Ayotte J, Peretz I, Hyde K (2002) Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain* 125(Pt 2):238–251.
31. Drayna D, Manichaikul A, de Lange M, Snieder H, Spector T (2001) Genetic correlates of musical pitch recognition in humans. *Science* 291(5510):1969–1972.
32. Peretz I, Cummings S, Dubé MP (2007) The genetics of congenital amusia (tone deafness): A family-aggregation study. *Am J Hum Genet* 81(3):582–588.
33. Hyde KL, et al. (2007) Cortical thickness in congenital amusia: When less is better than more. *J Neurosci* 27(47):13028–13032.
34. Loui P, Alsop D, Schlaug G (2009) Tone deafness: A new disconnection syndrome? *J Neurosci* 29(33):10215–10220.
35. Hyde KL, Zatorre RJ, Peretz I (2011) Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cereb Cortex* 21(2):292–299.
36. Foxton JM, Dean JL, Gee R, Peretz I, Griffiths TD (2004) Characterization of deficits in pitch perception underlying 'tone deafness' *Brain* 127(Pt 4):801–810.
37. Liu F, Patel AD, Fourcin A, Stewart L (2010) Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain* 133(Pt 6):1682–1693.
38. Peretz I, et al. (2002) Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron* 33(2):185–191.
39. Hyde KL, Peretz I (2004) Brains that are out of tune but in time. *Psychol Sci* 15(5): 356–360.
40. Vos PG, Troost JM (1989) Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Percept* 6:383–396.
41. Tufts JB, Molis MR, Leek MR (2005) Perception of dissonance by people with normal hearing and sensorineural hearing loss. *J Acoust Soc Am* 118(2):955–967.
42. Belin P, Fillion-Bilodeau S, Gosselin F (2008) The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behav Res Methods* 40(2):531–539.
43. Kohlrausch A, Fassel R, Dau T (2000) The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. *J Acoust Soc Am* 108 (2):723–734.
44. Pressnitzer D, McAdams S (1999) Two phase effects in roughness perception. *J Acoust Soc Am* 105(5):2773–2782.
45. Ritsma RJ (1967) Frequencies dominant in the perception of the pitch of complex sounds. *J Acoust Soc Am* 42(1):191–198.
46. Bidelman GM, Krishnan A (2009) Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *J Neurosci* 29(42):13165–13171.
47. Schwartz DA, Howe CQ, Purves D (2003) The statistical structure of human speech sounds predicts musical universals. *J Neurosci* 23(18):7160–7168.
48. Stewart L (2011) Characterizing congenital amusia. *Q J Exp Psychol (Hove)* 64(4): 625–638.
49. Tillmann B, Schulze K, Foxton JM (2009) Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds. *Brain Cogn* 71(3):259–264.
50. Williamson VJ, McDonald C, Deutsch D, Griffiths TD, Stewart L (2010) Faster decline of pitch memory over time in congenital amusia. *Adv Cogn Psychol* 6:15–22.
51. Gosselin N, Jolicoeur P, Peretz I (2009) Impaired memory for pitch in congenital amusia. *Ann N Y Acad Sci* 1169:270–272.
52. Bruckert L, et al. (2010) Vocal attractiveness increases by averaging. *Curr Biol* 20(2): 116–120.
53. Bregman AS (1990) *Auditory Scene Analysis* (MIT Press, Cambridge, MA).
54. Peretz I, Champod AS, Hyde K (2003) Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Ann N Y Acad Sci* 999:58–75.
55. Patterson RD, Allerhand MH, Giguère C (1995) Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform. *J Acoust Soc Am* 98(4):1890–1894.

Supporting Information

Cousineau et al. 10.1073/pnas.1207989109

SI Text

Audio Examples. The audio examples of stimuli include: (*i*) the voice minor second (m2) and perfect fifth (P5) chords of Exp. 1 depicted in Fig. 1; and (*ii*) one example for each of the harmonic (pure tone, harmonic-subset and octave-spaced), inharmonic (jittered, shifted, and 13st-spaced.), beating (diotic 0.75st), and nonbeating (dichotic 0.75st) stimuli of Exp. 3.

Roughness for Natural Stimuli Compared with Synthetic Complex Tones. Beating-based theories of consonance have argued that the amount of roughness found in dissonant chords is greater than that found in consonant chords, and ascribe the unpleasant sensation evoked by dissonant chords to the presence of roughness (1–3). One common feature of these classic studies is that they used chords composed of synthetic complex tones with equal amplitude harmonics. Such tones share the harmonic frequency relations of natural musical stimuli, but differ from them in a number of other ways (amplitude and phase of the partials, attack cues, etc.) and do not represent the variability that can be present in natural music listening situations. To investigate whether the relation between dissonance and roughness previously observed in synthetic tones was also present in the natural stimuli used here, we computed a Roughness Index for each of our stimuli, as well as for stimuli generated from synthetic tones. The stimuli were first passed through a bank of gamma-tone filters (4). The envelope of each filter output was extracted via the Hilbert transform, raised to the 0.3 power to simulate compression, and filtered between 20 and 200 Hz, using eighth-order Butterworth filters, to isolate the modulation rates that are standardly thought to contribute to roughness (5, 6). The Roughness Index was then calculated as the total power remaining in each band. Our synthetic tones consisted of the first six harmonics (with frequencies F0 to 6*F0) with equal amplitude and added in sine phase, as these stimuli have generally been used in models of dissonance based on beating (e.g., in refs. 1 and 3).

The variations of the Roughness Index obtained for chords composed of synthetic complex tones reproduce rather closely the ratings of such chords obtained from subjects in ref. 3. However, it is apparent that (*i*) the amount of beating physically present in the stimuli is generally larger for synthetic chords than for those generated from natural stimuli, and (*ii*) differences between intervals are more pronounced for the synthetic stimuli (Fig. S1). The voice stimuli in particular exhibit much less beating than do the other timbres, with little variation from interval to interval despite the large variation across intervals in perceived pleasantness.

To summarize the differences in beating between intervals heard as consonant and intervals heard as dissonant, we averaged the Roughness Index across the consonant and dissonant intervals that contributed to our behavioral measures of interval preference, for each of the three note timbres. The results shown in Fig. S2 illustrate that the roughness differences between consonant and dissonant intervals are indeed far more pronounced for intervals composed of synthetic tones than they are for those composed of naturally occurring instrument sounds.

In addition to the large differences in roughness between synthetic and natural intervals, there are pronounced differences even between saxophone and voice intervals. Consequently, if pleasantness ratings for chords were based even partly on roughness, we should see a difference between ratings for the two types of natural stimuli. Contrary to this prediction, normal-hearing listeners's ratings are quite similar for saxophone and voice intervals (Fig. S3). This finding is consistent with the idea that roughness has little bearing on their perception of chord pleasantness. For amusics however, ratings for voice intervals are consistently higher than ratings for saxophone intervals. One possible explanation for these results is that control subjects largely ignore roughness, having learned through a lifetime of exposure that roughness in music is not a reliable cue for dissonance. In contrast, amusics, lacking access to the harmonicity cue that we believe underlies dissonance, cannot learn that roughness is not a reliable cue, and thus show some sensitivity to roughness in their pleasantness judgments of chords. The difference in ratings between instrument type despite the lack of differentiation between different intervals likely reflects the fact that the roughness differences between intervals are small compared with the roughness differences between different timbres (Fig. S2).

Importantly, our stimuli were generated from real-world recordings of an instrument and a voice. It thus seems likely that the acoustic factors that influence consonance judgments under natural conditions are present in our stimuli. It is possible that for artificial stimuli with equal amplitude partials, which produce greater differences in roughness, we might have seen some degree of preference even in amusics. However, such stimuli are not representative of typical musical stimuli, in which variations in roughness are subtle and inconsistent.

We conclude from these analyses that roughness is not consistently present in dissonant chords, that roughness is sometimes present in consonant chords, and that roughness varies more as a function of instrumentation than as a function of whether a chord is consonant or dissonant. The pioneering work of the 1960s that established roughness as a cornerstone of musical aesthetics may thus have been a red herring.

Correlation Analyses of Preferences for Consonance and Preferences/Detection of Harmonicity and Beating. To further investigate individual differences in chord ratings and relate them to the perception of harmonicity and beating, we computed correlations between our measures of preference and detection. When the two groups of listeners were combined, the correlation between the preference for consonance and the harmonicity detection score was highly significant ($r = 0.76$, $P = 0.0001$), whereas the correlation between the preference for consonance and the beating detection score was not significant ($r = 0.20$, $P = 0.43$). These results are consistent with the results reported in ref. 7, in which preferences for consonance and harmonicity were correlated, whereas preference for consonance and beating were not. When computed within either the control group or the amusic group alone, however, none of these correlations were significant (Fig. S4, *Upper*). This finding is arguably unsurprising given the consistency within groups; in ref. 7 much larger cohorts were used, which overcame the modest individual differences present in normal listeners, most of whom have substantial preferences for consonance, harmonicity, and the absence of beating.

The very high scores obtained in both groups for beating detection mean that a ceiling effect could in principle be responsible for the absence of correlation between this measure and the preference for consonance. However, we also computed correlations between preference for consonance and our preference measures for harmonicity and beating; these correlations are more immune to this problem given that the preferences are not close to the maximum values possible given the rating scale. As expected, when the two groups were pooled together, a significant correlation was again observed between consonance and harmonicity ($r = 0.48$,

$P = 0.03$) but not between consonance and beating ($r = 0.17$, $P = 0.46$) (Fig. S4, *Lower*).

Dichotic Versions of the Harmonicity Stimuli. The frequency components of the harmonic stimuli were spaced more than a critical bandwidth apart in frequency, with the intention that the spectral modifications involved in creating the inharmonic stimuli would not alter the degree of beating. However, although beating is greatly reduced when frequency components are more than a critical band apart, there are conditions in which beats are nonetheless present, typically when frequency components are related by ratios that deviate slightly from small integer ratios (8). Dichotic presentation of frequency components is known to greatly reduce even these other forms of beats (9), and we therefore included alternate harmonicity test stimuli in which the even and odd numbered frequency components were played to opposite ears (dichotic presentation). No difference was found between the ratings obtained for diotic and dichotic versions of the harmonicity stimuli, and the ratings were thus combined in the analyses. The ratings and discrimination data for the diotic and dichotic versions of these stimuli are shown below in Figs. S5 and S6.

1. Plomp R (1967) Beats of mistuned consonances. *J Acoust Soc Am* 42(2):462–474.
2. Feeney MP (1997) Dichotic beats of mistuned consonances. *J Acoust Soc Am* 102(4): 2333–2342.
3. McDermott JH, Lehr AJ, Oxenham AJ (2010) Individual differences reveal the basis of consonance. *Curr Biol* 20(11):1035–1041.
4. Plomp R, Levelt WJ (1965) Tonal consonance and critical bandwidth. *J Acoust Soc Am* 38(4):548–560.
5. Kameoka A, Kuriyagawa M (1969) Consonance theory part I: Consonance of dyads. *J Acoust Soc Am* 45(6):1451–1459.
6. Kameoka A, Kuriyagawa M (1969) Consonance theory part II: Consonance of complex tones and its calculation method. *J Acoust Soc Am* 45(6):1460–1469.
7. Patterson RD, Allerhand MH, Giguère C (1995) Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform. *J Acoust Soc Am* 98(4):1890–1894.
8. Terhardt E (1974) Pitch, consonance, and harmony. *J Acoust Soc Am* 55(5):1061–1069.
9. Pressnitzer D, McAdams S (1999) Two phase effects in roughness perception. *J Acoust Soc Am* 105(5):2773–2782.

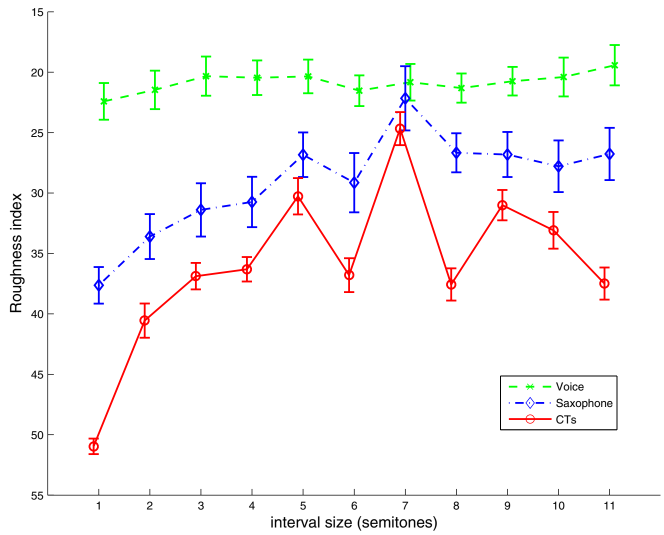

Fig. S1. Mean Roughness Index (averaged over eight different root notes) is plotted against interval size for chords composed of the three types of stimuli described in the text: recorded voice notes (green), recorded saxophone notes (blue), and synthetic complex tone notes (red). Note that the axis is inverted to make it congruent with the pleasantness rating axis in our figures, such that large amounts of roughness are lower on the graph. The variations of the Roughness Index for synthetic complex tone intervals closely reproduce behavioral ratings obtained in prior studies (e.g., ref. 6). For natural sounds however, the Roughness Index is in general smaller, and its variations more subtle, not resembling the rating pattern seen in behavioral data.

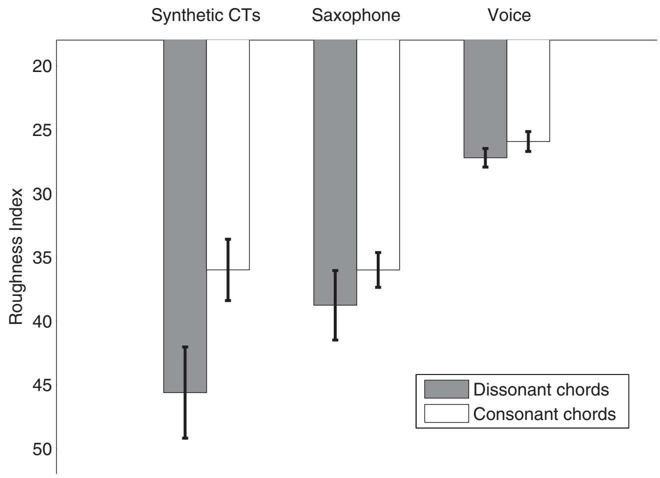

Fig. S2. Mean Roughness Index, averaged across the consonant and dissonant intervals that contributed to our behavioral measures of preference, for each of the three note timbres. Note that for consistency with Fig. S4, the roughness axis is again inverted here. Roughness differences between consonant and dissonant intervals are far more pronounced for intervals composed of synthetic complex tones than they are for those composed of naturally occurring instrument sounds. Large differences in overall roughness are also present between the different timbres of the synthetic and natural stimuli, and in fact exceed the difference between consonant and dissonant intervals within any particular timbre.

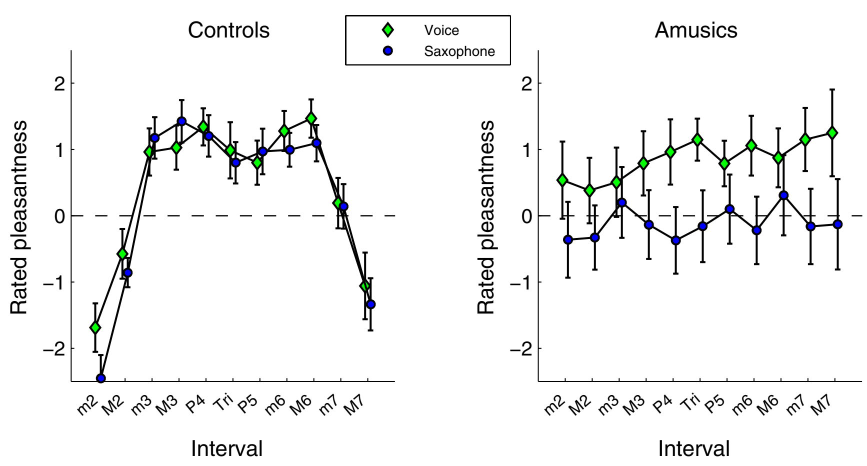

Fig. S3. Mean pleasantness ratings of chords by the control (*Left*) and amusic (*Right*) groups. For both groups, ratings are averaged over voice stimuli (green diamonds) on the one hand, and over saxophone stimuli (blue circles) on the other hand. Note that to reveal the difference between voice and saxophone in the amusics' ratings, the data in this figure is not *z*-scored.

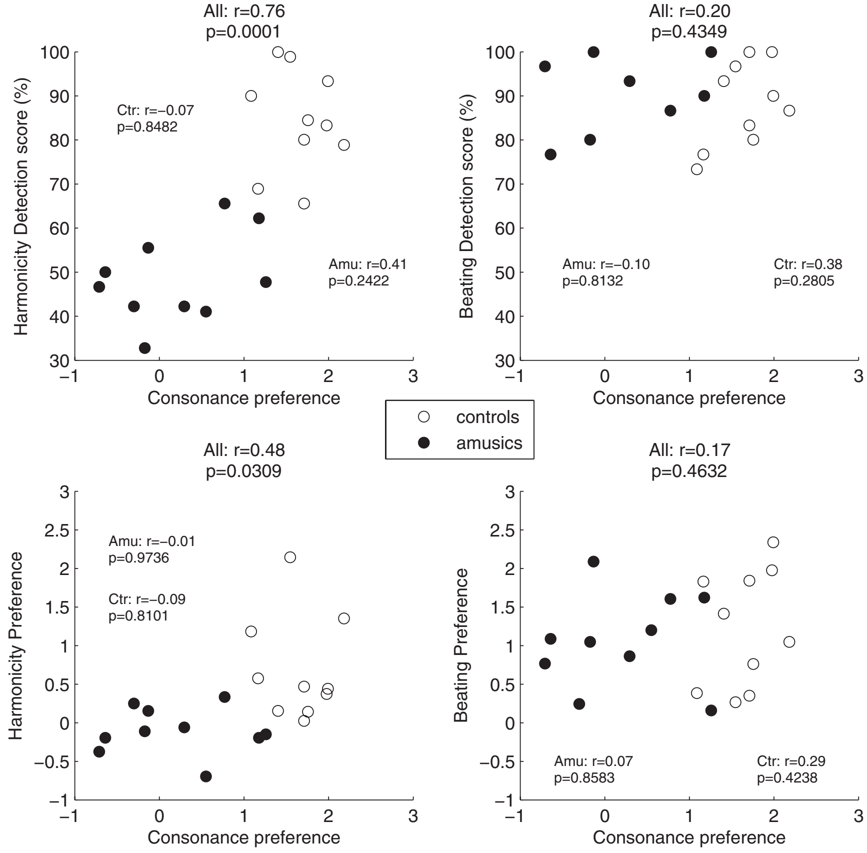

Fig. S4. Individual preferences for consonance plotted against detection score (*Upper*) or preference measure (*Lower*) for harmonicity (*Left*) and beating (*Right*). Correlations between the two scores are reported in each panel across groups of listeners (All), and for the amusic (amu) and the control (ctr) groups independently. Across groups, preference for consonance is correlated with both detection and preference for harmonicity, but with neither of those measures for beating.

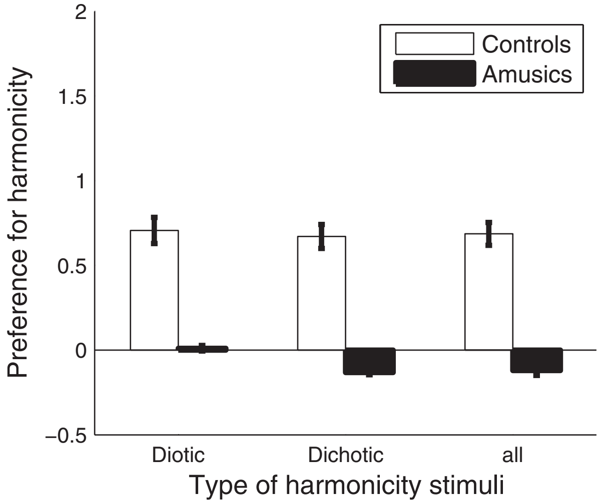

Fig. S5. Preference ratings for the diotic and dichotic versions of the harmonicity stimuli. Preferences averaged over both diotic and dichotic stimuli, as they appear in Fig. 5 (measure 1 of harmonicity preference), are also replotted here ("all"). The fact that the preference is not larger for diotic stimuli indicates that any residual beating contributed little to the ratings of the diotic stimuli. Here and elsewhere, error bars denote one SE about the mean.

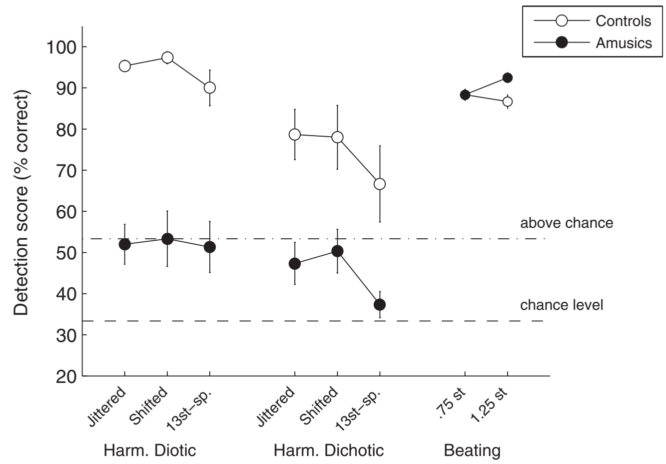

Fig. S6. Harmonicity and beating discrimination for each individual stimulus contrast used (three different types of inharmonicity × dichotic/diotic presentation for the harmonicity experiment, and two different frequency separations for the beating experiment). Simulations of a randomly responding individual performing 15 repetitions per condition of a 3-AFC task were used to derive a null distribution for task performance. Of the 10,000 obtained values, 95% were situated below the "significantly above chance" level depicted in this figure.

Table S1. Control and amusic group characteristics

Characteristics	Group		
	Amusics (n = 10)	Controls (n = 10)	P value of t test
Demographic characteristics			
Age (y)	65.6 ± 5.7	65.7 ± 4.2	0.97 (n.s.)
Sex	6 female	7 female	—
Education (y)	17.1 ± 3.6	16.4 ± 2.5	0.62 (n.s.)
Musical education (y)	1.1 ± 0.9	1.8 ± 1.5	0.23 (n.s.)
Audiogram (dB hearing loss)	20.2 ± 9.6	24.6 ± 13.0	0.40 (n.s.)
Music discrimination			
MBEA Melodic tests (% correct)	60.1 ± 8.4	91 ± 6.55	< 0.0001
MBEA Rhythmic test (% correct)	67.8 ± 9.0	89.8 ± 6.2	< 0.0001

The audiogram data are mean values of thresholds at 0.5, 1, 2, 3, 4, and 8 kHz in the two ears. The melodic test scores are expressed in percentages of correct responses and obtained on the scale, contour, and interval tests of the Montreal Battery of Evaluation of Amusia (MBEA) (1). The rhythmic test scores are obtained on the rhythm test of the same battery. Values are group means ± SD.

1. Peretz I, Champod AS, Hyde K (2003) Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Ann N Y Acad Sci* 999:58–75.

Other Supporting Information Files

Audio S1: Chord-Voice-m2 (WAV)
Audio S2: Chord-Voice-P5 (WAV)
Audio S3: Harm_PureTone (WAV)
Audio S4: Harm_Subset (WAV)
Audio S5: Harm_Octave-spaced (WAV)
Audio S6: Harm_Jittered (WAV)
Audio S7: Harm_Shifted (WAV)
Audio S8: Harm_13st-spaced (WAV)
Audio S9: Beat_Diotic (WAV)
Audio S10: Beat_Dichotic (WAV)